

University of Groningen

**Assessment of Volumetric versus Manual Measurement in Disseminated Testicular Cancer; No Difference in Assessment between Non-Radiologists and Genitourinary Radiologist**

Ozturk, Cigdem; Velleman, Ton; Bongaerts, Alphons H. H.; Bergman, L. M.; van Ginkel, Robert J.; Gietema, Jourik A.; Hoekstra, Harald J.

*Published in:*  
 PLoS ONE

*DOI:*  
 [10.1371/journal.pone.0168977](https://doi.org/10.1371/journal.pone.0168977)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
 Publisher's PDF, also known as Version of record

*Publication date:*  
 2017

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Ozturk, C., Velleman, T., Bongaerts, A. H. H., Bergman, L. M., van Ginkel, R. J., Gietema, J. A., & Hoekstra, H. J. (2017). Assessment of Volumetric versus Manual Measurement in Disseminated Testicular Cancer; No Difference in Assessment between Non-Radiologists and Genitourinary Radiologist. *PLoS ONE*, 12(1), [e0168977]. <https://doi.org/10.1371/journal.pone.0168977>

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

RESEARCH ARTICLE

# Assessment of Volumetric versus Manual Measurement in Disseminated Testicular Cancer; No Difference in Assessment between Non-Radiologists and Genitourinary Radiologist

Çiğdem Öztürk<sup>1☯</sup>, Ton Velleman<sup>1,2☯</sup>, Alphons H. H. Bongaerts<sup>2☯†</sup>, L. M. Bergman<sup>2</sup>, Robert J. van Ginkel<sup>1‡</sup>, Jourik A. Gietema<sup>3‡</sup>, Harald J. Hoekstra<sup>1‡\*</sup>

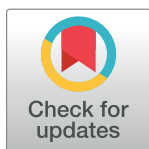
**1** Department of Surgical Oncology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands, **2** Radiology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands, **3** Medical Oncology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands

☯ These authors contributed equally to this work.

† Deceased.

‡ These authors also contributed equally to this work.

\* [h.j.hoekstra@umcg.nl](mailto:h.j.hoekstra@umcg.nl)



## OPEN ACCESS

**Citation:** Öztürk Ç, Velleman T, Bongaerts AHH, Bergman LM, van Ginkel RJ, Gietema JA, et al. (2017) Assessment of Volumetric versus Manual Measurement in Disseminated Testicular Cancer; No Difference in Assessment between Non-Radiologists and Genitourinary Radiologist. PLoS ONE 12(1): e0168977. doi:10.1371/journal.pone.0168977

**Editor:** Gayle E. Woloschak, Northwestern University Feinberg School of Medicine, UNITED STATES

**Received:** May 30, 2016

**Accepted:** December 11, 2016

**Published:** January 12, 2017

**Copyright:** © 2017 Öztürk et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are in the Supporting Information files.

**Funding:** The authors received no specific funding for this work.

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

### Background

The aim of this study was to assess the feasibility and reproducibility of semi-automatic volumetric measurement of retroperitoneal lymph node metastases in testicular cancer (TC) patients treated with chemotherapy versus the standardized manual measurements based on RECIST criteria.

### Methods

21 TC patients with retroperitoneal lymph node metastases of testicular cancer were studied with a CT scan of chest and abdomen before and after cisplatin based chemotherapy. Three readers, a surgical resident, a radiological technician and a radiologist, assessed tumor response independently using computerized volumetric analysis with Vitrea software® and manual measurement according to RECIST criteria (version 1.1). Intra- and inter-rater variability were evaluated with intra class correlations and Bland-Altman analysis.

### Results

Assessment of intra observer and inter observer variance proved non-significant in both measurement modalities. In particularly all intraclass correlation (ICC) values for the volumetric analysis were > .99 per observer and between observers. There was minimal bias in agreement for manual as well as volumetric analysis.

## Conclusion

In this study volumetric measurement using Vitrea software® appears to be a reliable, reproducible method to measure initial tumor volume of retroperitoneal lymph node metastases of testicular cancer after chemotherapy. Both measurement methods can be performed by experienced non-radiologists as well.

## Introduction

Diagnostic process for testicular cancer (TC) includes physical examination, laboratory tests for tumor markers alpha-fetoprotein (AFP), human chorionic gonadotropin (HCG), lactate dehydrogenase (LDH), and imaging tests, such as testicular ultrasound and computed tomography (CT) of abdomen and chest to assess regional and distant metastases. CT scans and tumor marker analysis make it possible to stage the extensiveness of disease and to classify the patient according to the International Germ Cell Consensus Classification (IGCCC) [1,2]. Staging results are fundamental in determining the prognosis and the optimal treatment strategy for each individual patient. Stage I non-seminomatous testicular germ cell tumor (NSTGCT) can be treated successfully with a wait-and-see policy, or unilateral nerve-sparing retroperitoneal lymph node dissection (RPLND), or one course of adjuvant cisplatin based chemotherapy after orchidectomy [2]. Disseminated disease is treated also with three or four courses of cisplatin-based chemotherapy. Resection of residual retroperitoneal tumor masses (RRRTM) after chemotherapy is an essential part of the combined therapy of NSTGCT [3,4]. CT of chest and abdomen serve as the main diagnostic tool to determine pre- and post-chemotherapy tumor deposit size, allowing evaluation of the response to treatment and guiding the decision to perform RRRTM, either conventional or via laparoscopic surgery [5]. An alternative approach to surgery can be observing patients after systemic chemotherapy, certainly if abdominal residual tumor masses have become undetectable (lesions < 1 cm) on CT scans and in some institution if the primary tumor did not contain teratoma components [6,7]. Residual metastatic disease after completion of chemotherapy can contain tumor necrosis, teratoma or viable tumor.

Optimal disease management is based on reliable and reproducible lesion measurement. Tumor response assessment has been standardized since the introduction of the bi-dimensional response criteria of the World Health Organization (WHO) which was more recently followed by one-dimensional (2D) response criteria based on RECIST (version 1.1), which is a set of rules defining overall tumor burden at baseline and objective tumor response and disease progression after systemic treatment [8,9].

In short, 2D measurements are obtained to assess the size of a tumor lesion. Measurements are obtained in the axial plane on a single slice, after careful selection from all axial planes in which the tumor lesion is visible. Assuming such single slice manual measurements in the axial plane comprise complete change in tumor size, neglects the fact that a tumor is an irregular three-dimensional lesion which is not a perfect sphere and therefore can be a misinterpretation of tumor burden. Additionally, in literature high inter-observer and intra-observer variability in manual tumor assessment are reported because of the difficulty to define the outer borders of the lesion to be measured and the variation in perception of the largest diameter [10,11].

An approach to improve reproducibility of lesion size measurements, and to reduce time and effort of measurements, combined with rapid technological developments, have led to the development of software allowing semi-automated lesion segmentation of tumor masses

[12,13]. Semi-automatic segmentation tools are utilized to acquire 3 dimensional (3D) images and accurate determination of tumor size alterations. Opposed to manual 2D measurements, in this 3D setting, asymmetrical tumor size changes can be detected and measured. Lymph node metastases are irregularly shaped and often closely located to adjacent tissue with similar contrast density. These characteristics, in combination with anatomical location and slice thickness, influence segmentation quality and increase measurement error, especially when measuring smaller shaped lesions [14,15]. The spread of TC to lymph nodes usually follows a predictable pattern with metastases occurring in retroperitoneal lymph nodes. Lymph node metastases in this region can be quite difficult to measure even with adequate contrast administrations and regular manual measurements, because of surrounding tissues with similar contrast densities causing indistinct borders of the tumor deposits [16]. Volumetric analysis is a new promising technique to measure a therapeutic response especially for lung and liver lesions, which can also be used to measure response of retroperitoneal lymph nodes of testicular cancer after chemotherapy [17].

The purpose of this pilot study was to assess the suitability and reproducibility of semi-automated volumetric analysis when applied to retroperitoneal lymph nodes in patients with disseminated TC, compared to 2D measurements based on RECIST between non-radiologists and a genitourinary radiologist.

## Materials and Methods

We consulted the Institutional Review Board (IRB) of the UMCG, and they confirmed that no formal written waiver for the need of ethics approval was required because of the retrospective design of the study. Also there was no written consent needed from the patients.

One of the supervising authors has had contact with some of the study participants as the surgeon who performed the resection of the lymph node metastases and therefore was a treating physician during a short period. This author did not perform the manual and volumetric measurements.

Authors did not collect any personally-identifying information from study participants. Furthermore, all data were detached of patient information prior to analysis.

## Study population

CT scans of Chest and Abdomen of 21 consecutive non-seminomatous testicular cancer patients (NSTGCT), treated at the University Medical Centre Groningen (UMCG), with a mean age of 36 (range 27–68) years, undergoing resection of residual retroperitoneal tumor mass (RRRTM) after cisplatin based combination chemotherapy were evaluated. Of the 21 patients, 16 patients (76%) had stage II disease, 2 patients (10%) had stage III disease and 3 patients (14%) had stage IV disease. According to the International Germ Cell Consensus (IGCC) classification the majority of patients, 15 (71%) had disease with good prognosis, 4 patients (19%) intermediate prognosis and 2 (10%) poor prognosis.

Lesions outside the retroperitoneal area were not considered for this analysis. All patients had pre-and post-chemotherapy CT imaging performed at the UMCG or the referring hospital.

Finally 21 patients were eligible with 28 individual lesions identified in the retroperitoneum.

## Data collection

All post-chemotherapy CT scans were performed at the UMCG with a multi-slice CT scanner (Sensation S64, Siemens Healthcare, Forchheim, Germany) according to the TC protocol, which encompasses fasting 4–6 hours prior to the scan after which scanning of abdomen and

thorax is performed with oral and intravenous contrast (100cc/flow 2.5). The reconstructed slice thickness for the standard protocol was 2.0 mm [18,19].

The majority of the pre-chemotherapy CT scans were performed in the referring hospitals. Depending on CT scanner quality and primary indication to perform a CT scan, reconstructed slice thickness was 2 mm for all scans.

All data were transferred to a dedicated advanced visualization workstation and were analyzed with a software package of Toshiba systems (Vitrea®/Vital Images, version 4) [20]. Vitrea® software is Vital Images' visualization software and commercially available for implementation. With this software 2D, 3D and 4D images are created of human anatomy from CT, MR (magnetic resonance) and PET (positron emission tomography) image data. The software provides options for cardiac, colon, vessel probe and other applications.

## Evaluation of retroperitoneal lymph node metastases

All manual and semi-automatic measurements of the labeled retroperitoneal lymph node metastases were executed independently by 2 non-radiologists (surgical resident and radiological technician) blinded for patient data and each others readings. These measurements were performed on a separate workstation without access to patient records. After transfer of CT scans to this separate workstation (without attachment of patient records), the readers measured the defined lesions while evaluating the CT images and performed the measurements. Tumor size was assessed based on RECIST criteria (1.1) [8,9].

The axial plane of the images chosen individually by each reader was used to draw long axis (LAD) and short axis (SAD) diameters in millimeters (Fig 1).

Semi-automatic evaluation of the retroperitoneal lymph nodes was also performed independently by the readers with Vitrea® software. This software package provided tools to perform volumetric measurements by semi-automatic segmentation on retroperitoneal lesions by manually selecting lesions in the axial and coronal plane followed by linear interpolation to create a 3D view. A nodule segmentation algorithm, based on radiodensity and grey scale differentiation, using Hounsfield unit values, is applied to separate the target lesions from background tissues and also morphological procedures are applied to separate these target lesions from adjacent tissue of the retroperitoneum. A visual assessment then followed on multi-planar reconstructions determining the contour of the lesion and in case of unsatisfactory segmentation manual adjustments could be made with correction tools if the lesion was not fully included in the segmentation field or if there was too much overlap of adjacent tissue (Fig 2). After satisfactory segmentation results, the software automatically generated and displayed a volume in cubic centimeters (cc).

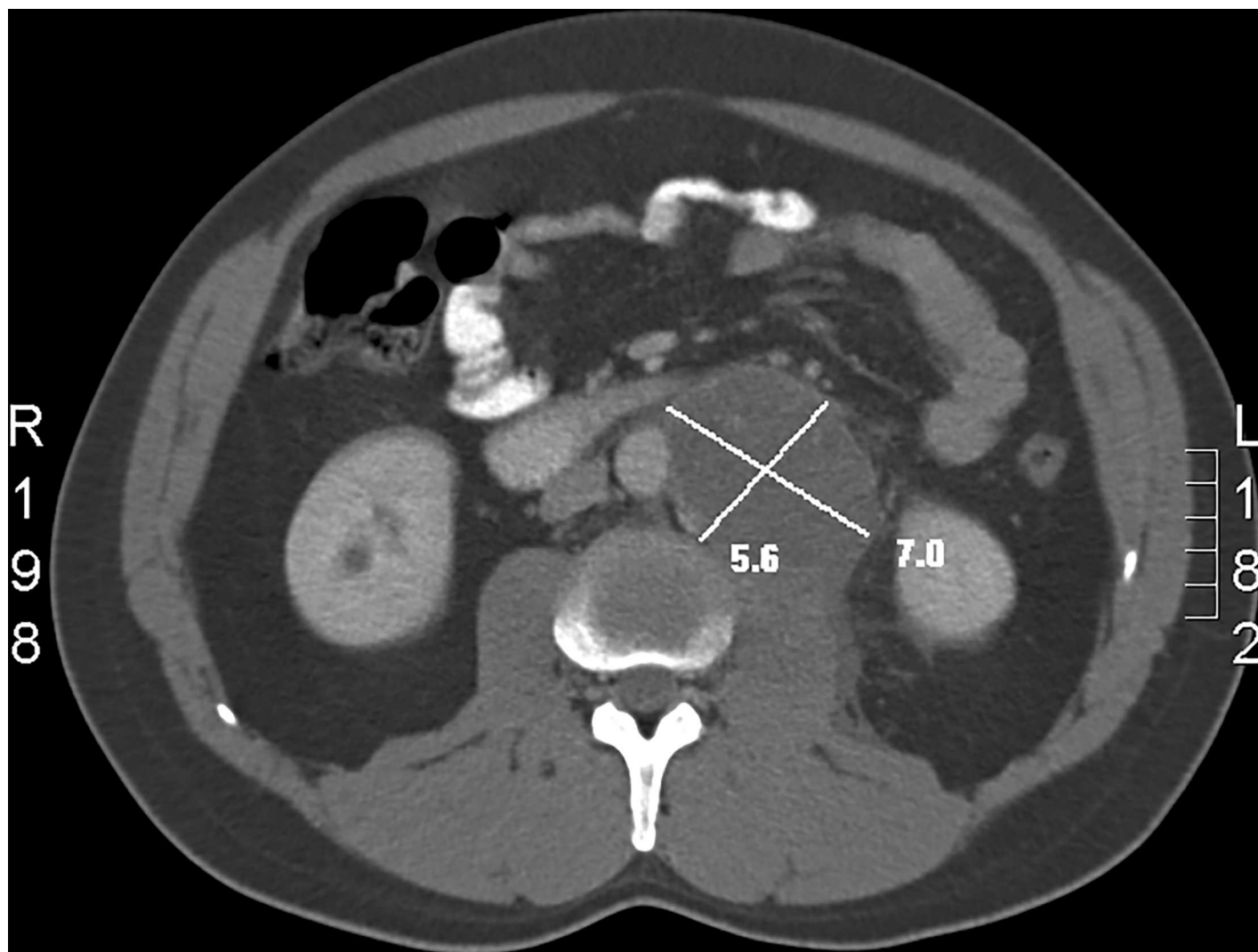
Finally, an experienced genitourinary radiologist also performed manual 2D and semi-automatic 3D measurements to make comparisons versus measurements made by the non-radiologists.

## Statistical analysis

Statistical analysis was performed with SPSS software version 18, City, US, Microsoft Excel, City, US.

Measurement meant either the tumor diameter in mm or volume in cc. In order to attain geometrically equivalent entities, the sphere volume formula  $V = \pi D^3 / 6$  was used to convert all volume measurements from  $\text{mm}^3$  to effective (spherical) diameters in mm.

Mean changes and range were calculated with SPSS. Descriptive statistics were used to summarize lesion characteristics and parameters (LAD, SAD, volume).



**Fig 1. Measurement according to RECIST in one dimension.** Long axis and short axis measurement in cm is shown in the figure.

doi:10.1371/journal.pone.0168977.g001

Intraclass correlation (ICC) values were used to assess intra-observer variation in the measurements according to RECIST and the volumetric method. Correlations were calculated to examine associations between the two measurement methods. Correlations with a coefficient  $<0.30$  were considered weak, between  $0.30$ – $0.50$  moderately strong, and  $>0.50$ , strong [21].

Intraclass correlation values were also applied to assess inter-rater variability of the two measuring methods. Furthermore, Bland-Altman analysis was performed to determine and visualize the inter-rater agreement between 2 readers. Bias (average difference between 2 readers), 95% confidence intervals and 95% limits of agreement were calculated.

## Results

### Radiologic characteristics of the retroperitoneal lymph nodes

Evaluation of CT scans of 21 patients, all diagnosed with NSTGCT, led to the identification of 28 individual retroperitoneal lymph nodes, which were analyzed manually and semi-automatically. Mean number of lymph nodes per patient was 1.5 (range 1–3). Distribution of lymph





**Fig 2. Vitrea® semi-automatic volumetric analysis.** Contour of the retroperitoneal lesion is marked.

doi:10.1371/journal.pone.0168977.g002

node size is presented in Table 1 and Fig 3. All these patients were in biochemical complete remission with normal tumor marker levels at time of the restaging CT-scan. Retroperitoneal lymph nodes which were resected either through a laparoscopic procedure or through a conventional laparotomy were localized as follows: 10 (35%) left para-aortic, 1 (4%) right para-aortic, 2 (7%) aortic bifurcation, 14 (50%) caudal of the left renal vein and 1 (4%) vena cava.

Table 2 provides an overview of the distribution of manual and volumetric measurement results as measured by 3 readers.

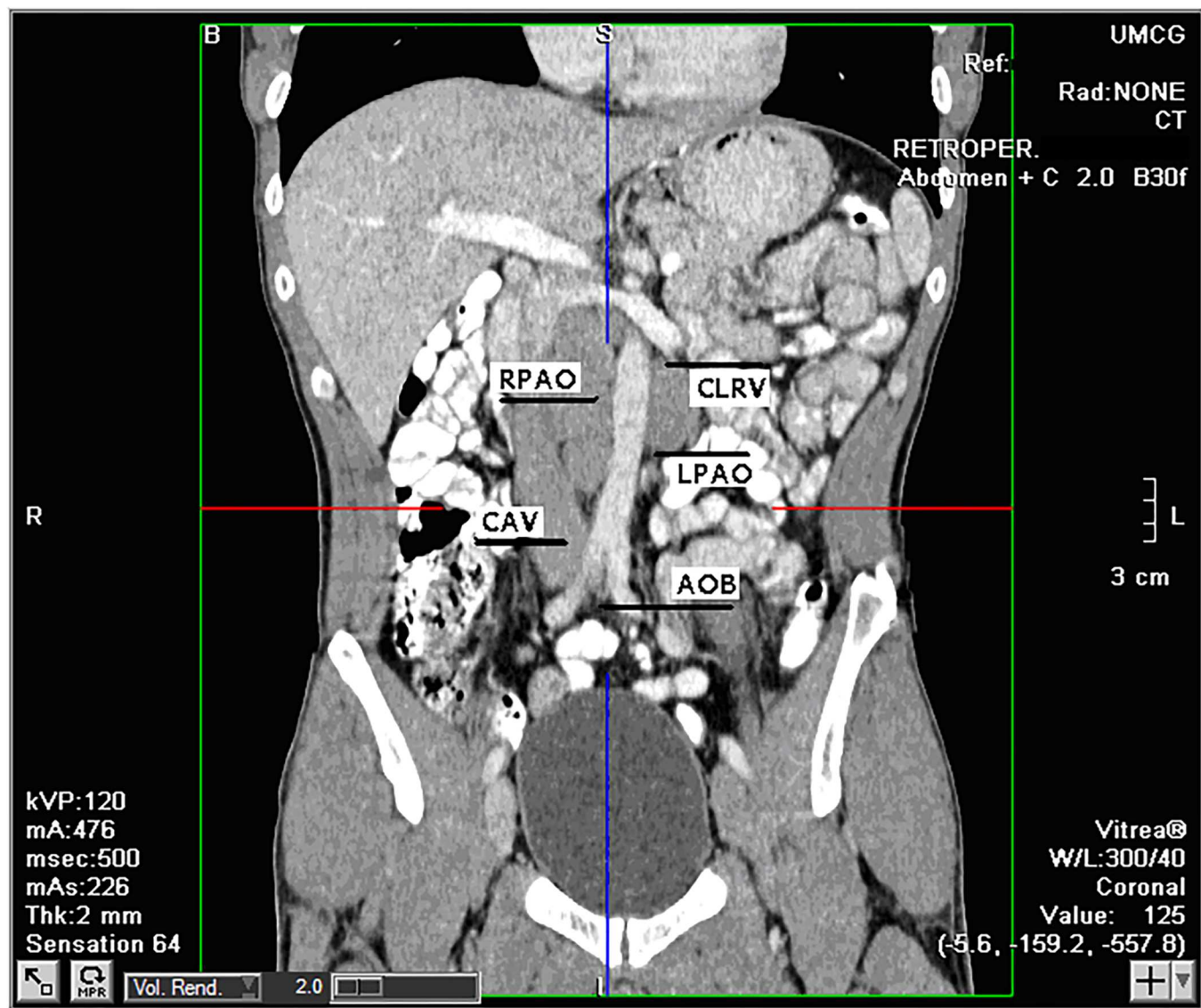
### Intra-observer variability: reproducibility by the same reader

Retroperitoneal lymph nodes were measured three times by 2 non-radiologist readers on the defined lesions. The ICC correlation for both readers was very significant with a correlation

**Table 1. Patient characteristics and lesion distribution.**

Patient characteristics	
Patients (n)	21
Age median (range)	31(22–63)
Lesions (n)	28
Stage II	16(76%)
Stage III	2(10%)
Stage IV	3(14%)
IGCC good	15(71%)
IGCC intermediate	4(19%)
IGCC poor	2(10%)
Anatomical site	
RPAO (n)	1
CLRV (n)	14
LPAO (n)	10
CAV (n)	1
AOB (n)	2

doi:10.1371/journal.pone.0168977.t001



**Fig 3. Distribution of retroperitoneal lymph nodes.** RPAO: right para-aortic, LPAO: left para-aortic, CLRV: caudal left renal vein, CAV: vena cava, AOB: aortic bifurcation.

doi:10.1371/journal.pone.0168977.g003

coefficient of .99 for reader 1 versus .98 for reader 2 with a corresponding P-value of  $< .001$  for both readers.

Volumetric analysis was repeated three times for each individual lymph node by the two readers showing an almost 100% concordance for the individual measurements in axial and coronal plane. All ICC values were  $> .99$  with a p-value of  $< .001$ .

### Inter-observer variability: consensus between readers

**Absolute measurements: manual and volumetric method compared.** Mean of the absolute measurements was used to calculate ICC values between two readers (Table 3). ICC values show an excellent correlation for both measurement methods between these readers with a p-value of  $< 0.001$ . Bland-Altman analysis (Fig 4) for testing the degree of agreement for manual as well as volumetric analysis shows minimal bias. Measurements of both methods of the two



**Table 2. Distribution of manual and volumetric lymph node measurements (mm).**

Parameters	Observer 1	Observer 2	Radiologist
<b>Pre chemo</b>			
LAD Mean $\pm$ SD	27.6 $\pm$ 14.2	28.1 $\pm$ 17.5	27.2 $\pm$ 13.5
LAD Median (range)	26.5 (8.2–76.4)	24.8 (7–91.3)	26.5 (5–70)
SAD Mean $\pm$ SD	24.3 $\pm$ 14.8	23.5 $\pm$ 13.5	23 $\pm$ 12
SAD Median (range)	21.8 (4–75.9)	20.7 (6–67)	21.5 (5–56)
Volume Mean $\pm$ SD	26.2 $\pm$ 13.1	25.7 $\pm$ 12.7	24.7 $\pm$ 12.9
Volume Median (range)	24.1 (8–70)	23.5 (8–68)	22 (7.8–69.8)
<b>Post chemo</b>			
LAD Mean $\pm$ SD	21.9 $\pm$ 14.4	21.3 $\pm$ 13.7	21.1 $\pm$ 14.3
LAD Median (range)	17.9 (7.5–64.1)	16.7 (7–60)	17 (6–60)
SAD Mean $\pm$ SD	18.9 $\pm$ 13.5	17.9 $\pm$ 13.6	17.8 $\pm$ 13.7
SAD Median (range)	14.6 (4.5–56.3)	13.5 (5.3–59)	13.5 (5–60)
Volume Mean $\pm$ SD	20.7 $\pm$ 11.9	20.5 $\pm$ 11.6	19.2 $\pm$ 11.7
Volume Median (range)	17.3 (6.9–51.7)	16.9 (7.1–51.4)	15.5 (6.7–50.4)

doi:10.1371/journal.pone.0168977.t002

readers were compared with measurements performed by an experienced radiologist. In Table 3 high ICC values are displayed. Results of the Bland Altman analysis are shown in Fig 4A–4F. In this graphical representation degree of agreement is displayed with 95% limits between observers on both manual and volumetric measurements. From these plots it was concluded that the spread between points are between the limits of agreement for both methods.

**Required time to perform volumetric analysis.** Time needed to perform volumetric measurements was calculated for 2 readers. Median measurement time for reader 1 was 1.42 (0.37–3.66) minutes and for reader 2 was 2.69 (range 1.19–4.54) minutes. Time of lymph node selection and transfer to another workstation where Vitrea® analysis could be performed was approximately 10–15 minutes. Time to perform RECIST was not measured. Characteristics of the measurements of the radiologist are lacking.

## Discussion

Measuring tumor burden and tumor response to therapy as accurate as possible in oncological practice remains a challenge. Limitations of manual measurements according to RECIST have been described in literature. An important limitation is that tumor measurement occurs at differently chosen slides by the readers, leading to a large intra- and inter-observer variation [10,11]. Volumetric images have gained importance in medical applications in recent years, resulting in software products allowing volumetric analysis. In liver surgery volumetry is important to calculate postoperative liver volume and liver function resulting in safety enhancement of liver resection procedures [12,22,23].

**Table 3. Overview intra class correlation values between observers across methods.**

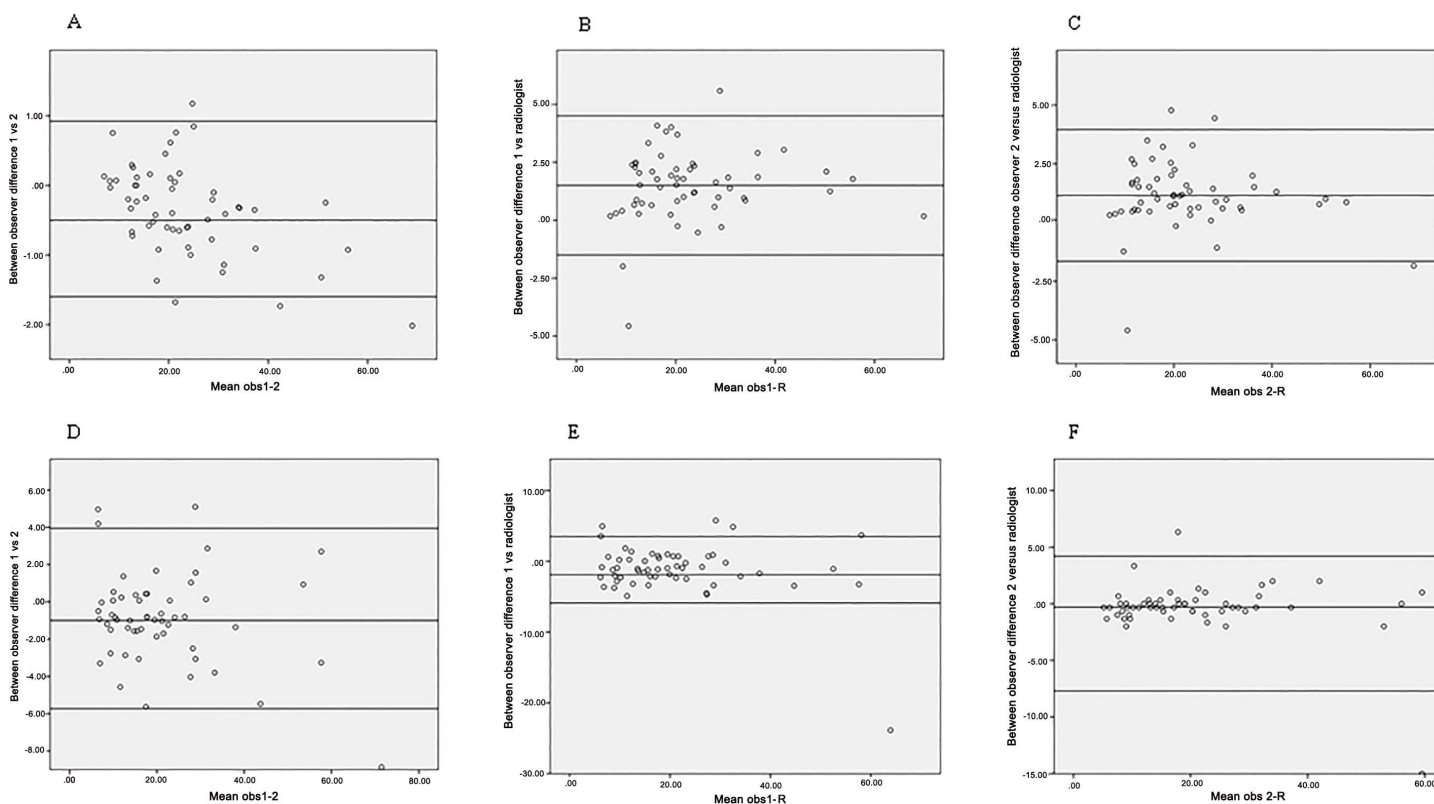
Volumetric	ICC	P-value
Observer 1–2	.99	.000
Observer 1-radiologist	.99	.000
Observer 2- radiologist	.99	.000
<b>Manual</b>		
Observer 1–2	.97	.000
Observer 1-radiologist	.96	.000
Observer 2-radiologist	.98	.000

doi:10.1371/journal.pone.0168977.t003

In the present study, radiological tumor response after chemotherapy in patients with disseminated non-seminomatous TC with retroperitoneal lymph node metastases was evaluated using two measurement procedures, namely 2D measurements based on the RECIST criteria and semi-automatic volumetric analysis with specialized computer software.

Results show a high intra-observer agreement for the measured tumor volumes with the used software package. Measurement of the retroperitoneal lymph node volumes performed by each reader, both non-radiologists, showed also high correlations. This corresponds with findings of semi-automatic volumetric analysis in lung nodules which are promising and show reproducible measurements with high accuracy [24–26]. In the present study intra-observer agreement for the 2D manual measurements was also high with ICC values of .99. In a previous study concerning intra-observer variability in the response evaluation according to RECIST in solid tumors such as breast cancer and colorectal cancer, correlation values ranged between 0.76–0.96 [27]. Intra-observer agreement with these manual measurements in the current study was higher than expected based on previous literature on this subject.

In the present study, user experience was assessed by measuring inter-observer agreement. Measurements were performed by non-radiologists as well as a radiologist to evaluate its user friendliness and reliable application in the routine hospital. Inter-observer variability for semi-automatic measurements was comparable for manual measurements, with acceptable limits of agreement. When we looked at absolute measurement values both measurement modalities were highly reliable. In another study, inter-observer variability was significantly lower in the semi-automatic method [28].



**Fig 4. Bland Altman plots.** (A) Display Bland Altman plot observer 1 and 2 volumetric method (cc). (B) Display Bland Altman plot observer 1 and R (radiologist) volumetric method (cc). (C) Display Bland Altman plot observer 2 and R volumetric method (cc). (D) Display Bland Altman plot observer 1 and 2 manual method (mm). (E) Display Bland Altman plot observer 1 and R manual method (mm). (F) Display Bland Altman plot observer 2 and R manual method (mm). Mean difference and limits of agreement are displayed as reference lines.

doi:10.1371/journal.pone.0168977.g004

Good results were obtained with volumetric analysis with a high level of agreement between observers and methods. It should be taken into account that the complexity of the parameters still require a lot of work to produce an informative and reliable image. Also, volumetric analysis of lymph node metastases compared to, for example, lung nodules can be more difficult since lymph node metastases are often irregularly shaped with inhomogeneous contrast enhancement and are in the proximity of similar density tissues.

Tumor volumetry can be time consuming because of the segmentation process. However in experienced hands of, a radiologist, a radiological technician or a resident the time investment seems appropriate. Also, results from this study showed that novices with some radiological experience and an experienced assessor such as a radiologist were both able to reliably predict tumor volume using the semi-automatic computer software. This validation study has been conducted in a center of expertise for patients with testicular cancer for over more than three decades. In present times there is a paradigm shift in the radiologic evaluation of lesions. Where in the past the executing surgeon relied 100% on the readings of a radiologist, nowadays the surgeon and a surgical team need tools to get a quick and reliable impression of lesions before performing a specific surgery. We have found that the volumetric method can be performed by an experienced non-radiologist as well and that the required time to obtain and measure a reliable 3D image is under 3 minutes. Such a study for volumetric measurements is lacking in literature for retroperitoneal lymph node metastases in testicular cancer. The new technology facilitates the surgical-oncologist or the uro-oncologist in the conventional or laparoscopic surgical treatment planning of the resection of tumor masses.

In conclusion, the pilot study showed that radiological technicians and (surgical/urology) residents could perform volumetric and manual measurements in a reliable manner with a high intra and inter-observer reliability. With the increase in the number of cancer patients receiving chemotherapy, automatic and semiautomatic segmentation tools will likely play an important role in the evaluation of tumor response to treatment, e.g. a cost-effective way to measure tumor response and prepare the surgical or uro-oncologist in the resection of retroperitoneal lymph node metastases. For now, clinical guidelines based on manual measurement in 2D in the axial plane according to RECIST criteria remain the standard method for assessing a radiological tumor response to anti-cancer therapy. This study has shown that the volumetric method is reproducible, user friendly and is promising as a more accurate method in the surgical field since it provides a 3D view. Especially the latter can make a clinical difference with more detailed surgical planning and can contribute to achieving better outcomes in patient care. Since tumor response criteria for volumetric measurements are not yet validated, combining the clinical outcome in a larger prospective study group could provide standardized volumetric measurements.

## Supporting Information

**S1 Fig. Measurement according to RECIST in one dimension.** Long axis and short axis measurement in cm is shown in the figure.

(DOCX)

**S2 Fig. Vitrea® semi-automatic volumetric analysis.** Contour of the retroperitoneal lesion is marked.

(DOCX)

**S3 Fig. Distribution of retroperitoneal lymph nodes.** RPAO: right para-aortic, LPAO: left para-aortic, CLRV: caudal left renal vein, CAV: vena cava, AOB: aortic bifurcation.

(DOCX)

**S4 Fig. Bland Altman plots.** (A) Display Bland Altman plot observer 1 and 2 volumetric method (cc). (B) Display Bland Altman plot observer 1 and R (radiologist) volumetric method (cc). (C) Display Bland Altman plot observer 2 and R volumetric method (cc). (D) Display Bland Altman plot observer 1 and 2 manual method (mm). (E) Display Bland Altman plot observer 1 and R manual method (mm). (F) Display Bland Altman plot observer 2 and R manual method (mm). Mean difference and limits of agreement are displayed as reference lines. (DOCX)

**S1 Table. Patient characteristics and lesion distribution.** RPAO: right para-aortic, CLRV: caudal left renal vein, LPAO: left para-aortic, CAV: vena cava, AOB: aortic bifurcation. (DOCX)

**S2 Table. Distribution of manual and volumetric lymph node measurements (mm).** (DOCX)

**S3 Table. Overview intra class correlation values between observers across methods.** (DOCX)

## Acknowledgments

A.H.H. Bongaerts MD, genitourinary radiologist, deceased recently and the authors appreciate the unselfish assistance in finalizing this manuscript by his colleagues L.M.Bergman, radiologist UMCG and P.M.A. van Ooijen, MSc, PhD, CPHIT, Associate Professor of Medical Imaging Informatics UMCG.

## Author Contributions

**Conceptualization:** ÇÖ TV HJH.

**Data curation:** ÇÖ TV AHHB.

**Formal analysis:** ÇÖ TV.

**Investigation:** ÇÖ TV AHHB.

**Methodology:** ÇÖ TV AHHB.

**Project administration:** ÇÖ TV JAG HJH.

**Resources:** ÇÖ TV AHHB.

**Supervision:** ÇÖ JAG RJG HJH.

**Validation:** ÇÖ TV AHHB LMB.

**Visualization:** ÇÖ TV RJG HJH.

**Writing – original draft:** ÇÖ TV.

**Writing – review & editing:** ÇÖ TV LMB JAG RJG HJH.

## References

1. International Germ Cell Consensus Classification: a prognostic factor-based staging system for metastatic germ cell cancers. International Germ Cell Cancer Collaborative Group. *J Clin Oncol.* 1997; 15 (2):594–603. doi: [10.1200/jco.1997.15.2.594](https://doi.org/10.1200/jco.1997.15.2.594) PMID: [9053482](https://pubmed.ncbi.nlm.nih.gov/9053482/)
2. Schmoll HJ, Jordan K, Huddart R, Pes MP, Horwich A, Fizazi K et al. Testicular non-seminoma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2010; Suppl 5:147–54.

3. Oldenburg J, Alfsen GC, Lien HH, Aass N, Waehre H, Fossa SD. Postchemotherapy retroperitoneal surgery remains necessary in patients with nonseminomatous testicular cancer and minimal residual tumor masses. *J Clin Oncol*. 2003; 21(17):3310–7. doi: [10.1200/JCO.2003.03.184](https://doi.org/10.1200/JCO.2003.03.184) PMID: [12947067](https://pubmed.ncbi.nlm.nih.gov/12947067/)
4. Kesler KA, Kruter LE, Perkins SM, Rieger KM, Sullivan KJ, Runyan ML et al. Survival after resection for metastatic testicular nonseminomatous germ cell cancer to the lung or mediastinum. *Ann Thorac Surg*. 2011; 91(4):1085–93. doi: [10.1016/j.athoracsur.2010.12.034](https://doi.org/10.1016/j.athoracsur.2010.12.034) PMID: [21440128](https://pubmed.ncbi.nlm.nih.gov/21440128/)
5. Oztürk C, van Ginkel RJ, Krol RM, Gietema JA, Hofker HS, Hoekstra HJ. Laparoscopic resection of a residual retroperitoneal tumor mass of nonseminomatous testicular germ cell tumors. *Surg Endosc*. 2012; 26(2):458–67. doi: [10.1007/s00464-011-1901-0](https://doi.org/10.1007/s00464-011-1901-0) PMID: [21938584](https://pubmed.ncbi.nlm.nih.gov/21938584/)
6. Heidenreich A<sup>1</sup>, Knüchel-Clarke R, Pfister D. Importance of pathology for Therapy planning of testicular germ cell tumors. *Pathologe*. 2014; 35(3):266–73. doi: [10.1007/s00292-014-1903-5](https://doi.org/10.1007/s00292-014-1903-5) PMID: [24771259](https://pubmed.ncbi.nlm.nih.gov/24771259/)
7. Debono DJ, Heilman DK, Einhorn LH, Donahue JP. Decision analysis for avoiding postchemotherapy surgery in patients with disseminated nonseminomatous germ cell tumors. *J Clin Oncol*. 1997; 15:1455–64. doi: [10.1200/jco.1997.15.4.1455](https://doi.org/10.1200/jco.1997.15.4.1455) PMID: [9193340](https://pubmed.ncbi.nlm.nih.gov/9193340/)
8. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J of Cancer*. 2009; 45(2):228–47.
9. Tirkes T, Hollar MA, Tann M, Kohli MD, Akisik F, Sandrasegaran K. Response criteria in oncologic imaging: review of traditional and new criteria. *Radiographics*. 2013; 33(5):1323–41. doi: [10.1148/rg.335125214](https://doi.org/10.1148/rg.335125214) PMID: [24025927](https://pubmed.ncbi.nlm.nih.gov/24025927/)
10. Marten K, Auer F, Schmidt S, Kohl G, Rummeny EJ, Engelke C. Inadequacy of manual measurements compared to automated CT volumetry in assessment of treatment response of pulmonary metastases using RECIST criteria. *Eur Radiol*. 2006; 16(4):781–90. doi: [10.1007/s00330-005-0036-x](https://doi.org/10.1007/s00330-005-0036-x) PMID: [16331462](https://pubmed.ncbi.nlm.nih.gov/16331462/)
11. Erasmus JJ, Gladish GW, Broemeling L, Sabloff BS, Truong MT, Herbst RS et al. Interobserver and intraobserver variability in measurement of non-small-cell carcinoma lung lesions: implications for assessment of tumor response. *J Clin Oncol*. 2003; 21:2574–82. doi: [10.1200/JCO.2003.01.144](https://doi.org/10.1200/JCO.2003.01.144) PMID: [12829678](https://pubmed.ncbi.nlm.nih.gov/12829678/)
12. Kalkmann J, Ladd SC, de Greiff A, Forsting M, Statta J. Suitability of semi-automated tumor response assessment of liver metastases using a dedicated software package. *Rofo*. 2012; 184(9):795–804.
13. Höink AJ<sup>1</sup>, Weßling J, Koch R, Schülke C, Kohlhaase N, Wassenaar L et al. Comparison of manual and semi-automatic measuring techniques in MSCT scans of patients with lymphoma: a multicentre study. *Eur Radiol*. 2014; 24(11):2709–18. doi: [10.1007/s00330-014-3283-x](https://doi.org/10.1007/s00330-014-3283-x) PMID: [25192795](https://pubmed.ncbi.nlm.nih.gov/25192795/)
14. Petrou M, Quint LE, Nan B, Baker LH. Pulmonary nodule volumetric measurement variability as a function of CT slice thickness and nodule morphology. *AJR Am J Roentgenol*. 2007; 188:306–12. doi: [10.2214/AJR.05.1063](https://doi.org/10.2214/AJR.05.1063) PMID: [17242235](https://pubmed.ncbi.nlm.nih.gov/17242235/)
15. Reeves AP, Chan AB, Yankelevitz DF, Henschke CI, Kressler B, Kostis WJ. On measuring the change in size of pulmonary nodules. *IEEE Trans Med Imaging*. 2006; 25:435–50. doi: [10.1109/TMI.2006.871548](https://doi.org/10.1109/TMI.2006.871548) PMID: [16608059](https://pubmed.ncbi.nlm.nih.gov/16608059/)
16. Hopper KD, Kasales CJ, Van Slyke MA, Schwartz TA, TenHave TR, Jozefiak JA. Analysis of interobserver and intraobserver variability in CT tumor measurements. *Am J Roentgenol*. 1996; 167:851–54.
17. Basilio C, Villeda C, Culebro C, Rodríguez-Covarrubias F, Castillejos-Molina R. Volumetric assessment of lymph node metastases in patients with non-seminomatous germ cell tumours treated with chemotherapy. *Can Urol Assoc J*. 2015; 9 (5–6): 247–51.
18. Buerke B, Puesken M, Beyer F, Gerres J, Weckesser M, Seifarth H, et al. Semiautomatic lymph node segmentation in multislice computed tomography: impact of slice thickness on segmentation quality, measurement precision, and interobserver variability. *Invest Radiol*. 2010; 45:82–88. doi: [10.1097/RLI.0b013e3181c443e1](https://doi.org/10.1097/RLI.0b013e3181c443e1) PMID: [20027120](https://pubmed.ncbi.nlm.nih.gov/20027120/)
19. Fabel M, Biederer J, Jochens A, Bornemann L, Soza G, Heller M, et al. Semi-automated volumetric analysis of artificial lymph nodes in a phantom study. *Eur J Radiol*. 2011; 80(3):451–7.
20. <https://www.vitalimages.com/solutions/clinical-insight/vitrea-software>.
21. Cohen J. Statistical power analysis for the behavioural sciences. New York: Academy Press: New York, 1988.
22. Kishi Y, Abdalla EK, Chun YS, Zorzi D, Madoff DC, Wallace MJ et al. Three hundred and one consecutive extended right hepatectomies: evaluation of outcome based on systematic liver volumetry. *Ann Surg*. 2009; 250:540–8. doi: [10.1097/SLA.0b013e3181b674df](https://doi.org/10.1097/SLA.0b013e3181b674df) PMID: [19730239](https://pubmed.ncbi.nlm.nih.gov/19730239/)
23. Lemke AJ, Brinkmann MJ, Schott T, Niehues SM, Settmacher U, Neuhaus P et al. Living donor right liver lobes: preoperative CT volumetric measurement for calculation of intraoperative weight and volume. *Radiology*. 2006; 240:736–42. doi: [10.1148/radiol.2403042062](https://doi.org/10.1148/radiol.2403042062) PMID: [16868277](https://pubmed.ncbi.nlm.nih.gov/16868277/)



24. Marten K, Auer F, Schmidt S, Kohl G, Rummeny EJ, Engelke C. Inadequacy of manual measurements compared to automated CT volumetry in assessment of treatment response of pulmonary metastases using RECIST criteria. *European Radiology*. 2006; 16:781–90. doi: [10.1007/s00330-005-0036-x](https://doi.org/10.1007/s00330-005-0036-x) PMID: [16331462](https://pubmed.ncbi.nlm.nih.gov/16331462/)
25. Bolte H, Riedel C, Jahnke T, Inan N, Freitag S, Kohl G et al. Reproducibility of computer-aided volumetry of artificial small pulmonary nodules in ex vivo porcine lungs. *Invest Radiol*. 2006; 41:28–35 PMID: [16355037](https://pubmed.ncbi.nlm.nih.gov/16355037/)
26. Scholten ET<sup>1</sup>, de Jong PA, Jacobs C, van Ginneken B, van Riel S, Willemink MJ et al. Interscan variation of semi-automated volumetry of subsolid pulmonary nodules. *Eur Radiol*. 2015; 25(4):1040–7. doi: [10.1007/s00330-014-3478-1](https://doi.org/10.1007/s00330-014-3478-1) PMID: [25413965](https://pubmed.ncbi.nlm.nih.gov/25413965/)
27. Suzuki C, Torkzad MR, Jacobsson H, Aström G, Sundin A, Hatschek T et al. Interobserver and intraobserver variability in the response evaluation of cancer therapy according to RECIST and WHO-criteria. *Acta Oncol*. 2010; 49:509–14. doi: [10.3109/02841861003705794](https://doi.org/10.3109/02841861003705794) PMID: [20397778](https://pubmed.ncbi.nlm.nih.gov/20397778/)
28. Buerke B, Puesken M, Mütter S, Weckesser M, Gerss J, Heindel W et al. Measurement accuracy and reproducibility of semiautomated metric and volumetric lymph node analysis in MDCT. *AJR Am J Roentgenol*. 2010; 195:979–85. doi: [10.2214/AJR.10.4010](https://doi.org/10.2214/AJR.10.4010) PMID: [20858828](https://pubmed.ncbi.nlm.nih.gov/20858828/)